心理科学进展 2020, Vol. 28, No. 6, 883–892
Advances in Psychological Science DOI: 10.3724/SP.J.1042.2020.00883

·研究构想(Conceptual Framework)·

音乐句法加工的认知机制与音乐结构的影响模式*

张晶晶[1] 梁啸岳[2] 陈伊笛[2] 陈庆荣[1]

([1]南京师范大学心理学院; [2]南京师范大学音乐学院, 南京 210097)

摘 要 音乐和语言是人类最重要的两种交流系统。与语言一样, 音符的排列和组织也是建立在一定的句法规则之上。尽管现有研究发现听众具有感知音乐句法的能力, 音乐句法加工的认知机制以及影响因素仍不清楚。基于此, 拟深入探究预期和整合在音乐音高句法加工中的作用, 以及音乐层级结构和时间结构对音高句法加工的影响。以期进一步揭示音乐句法加工的本质, 为音乐和语言的比较以及探索人类更一般的交流机制提供实证依据。

关键词 音乐句法; 层级结构; 时间结构; 认知机制; 神经机制

分类号 B842

音乐和语言是人类最重要的两种交流方式。和语言一样, 音符之间的排列并不是随意组织, 而是遵循一定的句法规则(Lerdahl & Jackendoff, 1983; 叶铮, 周晓林, 2006; 周临舒, 蒋存梅, 杨玉芳, 2012)。建立在调性音乐基础之上的音高组织规则是一种突出的句法现象, 既约束着和弦的构成与连接, 也体现在旋律音高的进行之中, 从而构筑起整个乐曲的进行和发展(Patel, 2010; 马谐, 杨玉芳, 张秋月, 2016; 张晶晶, 杨玉芳, 2017)。

音乐句法在音乐感知中起着重要作用, 制约着听众对音乐事件的期待, 并在此基础上促成音乐情感和美学体验的产生(Koelsch, 2014)。现有音乐句法研究主要集中于考察人们是否具有音乐句法的内隐知识。实验结果发现, 由于长期暴露在调性音乐环境中, 听众普遍具有感知音乐句法的能力(Koelsch, Vuust, & Friston, 2019)。进一步的科学问题是, 音乐句法加工的认知机制是什么? 音乐句法加工如何受到音乐结构的调节?

1 国内外研究现状

1.1 音乐句法加工的认知机制

已有研究主要采用行为和认知神经科学的方法考察人们能否加工音乐句法。实验中, 研究者通常给被试呈现一组较短的和弦序列, 通过操纵终止和弦(目标和弦)或者先前的背景和弦, 构建句法合理和句法不合理的和弦序列, 从而考察人们的音乐句法加工能力。行为研究发现, 当和弦序列句法不合理时, 相对句法合理, 听众对目标和弦的加工速度更慢(Bharucha & Stoeckig, 1987; Tillmann, Bigand, & Pineau, 1998), 且句法违反程度非常微弱时, 这种差异仍然存在(Bigand & Pineau, 1997; Bigand, Tillmann, Poulin, D'Adamo, & Madurell, 2001)。ERP 结果显示, 句法不合理的目标和弦在 100~350 ms 处诱发右前负波(early right anterior negativity, ERAN) (Jentschke, Friederici, & Koelsch, 2014; Koelsch, Gunter, Friederici, & Schröger, 2000; Koelsch, Jentschke, Sammler, & Mietchen, 2007), 且在晚期阶段诱发 N5 (Koelsch et al., 2000; Poulin-Charronnat, Bigand, & Koelsch, 2006)或者 LPC (Lagrois, Peretz, & Zendel, 2018; Patel, Gibson, Ratner, Besson, & Holcomb, 1998)。ERAN 反映早期对句法违反的探测, N5 或者 LPC

收稿日期: 2019-09-03

* 国家自然科学基金青年项目(31800914), 江苏省社会科学基金项目研究成果(19YYC003), 国家社会科学基金重点项目(18AYY010), 江苏省高校自然科学研究面上项目(18KJD190003)。

通信作者: 陈庆荣, E-mail: jscqr80@sina.com

反映了随后句法整合过程和对音乐结构的再分析(Koelsch et al., 2000; Patel et al., 1998)。MEG (Maess, Koelsch, Gunter, & Friederici, 2001)和fMRI (Tillmann, Janata, & Bharucha, 2003)研究进一步定位了音乐句法加工脑区，发现音乐句法违反相比句法合理，在布洛卡区以及右侧镜像区域出现了更强的激活。

上述一系列研究表明，听众能够感知音乐句法，音乐和声背景提供的信息即时影响了新出现和弦的加工。那么，潜在的重要问题是，音乐句法的加工是怎样进行的？我们认为存在两种可能的认知机制。第一种是基于预期的观点：随着音乐的展开，听众根据句法知识对接下来将要出现的和弦建立了预期，当目标和弦与预期不一致时，听众感到诧异，导致对目标和弦加工更困难。第二种是基于整合的观点：当目标和弦与音乐背景句法不合理时，在目标和弦出现之后，听众需要花费更多的认知资源，将目标和弦整合到音乐背景的表征之中。

在语言加工领域，对于整合和预期也曾有过长久的争论(Kuperberg & Jaeger, 2016; van Petten & Luka, 2012)。在早期阶段，预期被严重忽视，整合观点占据主导地位。20 世纪 80 年代，研究者们认为句子理解是缓慢进行的，只有当新词语呈现结束并且完成词汇理解之后，新词语与句子背景的整合才开始进行(Kintsch, 1988)。到 20 世纪 90 年代末，研究者们虽然发现整合不是延迟而是即时发生的，但是也只有当新词语出现后句子背景才会对其加工产生影响(Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995)。近些年，随着实验范式和实验技术的进步，预期在语言加工中的作用开始凸显出来(e.g. Chen et al., 2016; DeLong, Urbach, & Kutas, 2005; Ito, Corley, Pickering, Martin, & Nieuwland, 2016; Ito, Pickering, & Corley, 2018; Li, Zhang, Xia, & Swaab, 2017; Maess, Mamashli, Obleser, Helle, & Friederici, 2016)。一系列研究支持了句法预期在语言加工中的作用，例如，Lau, Stroud, Plesch 和 Phillips (2006)给被试呈现两类句子，比如句(1)和句(2)，这两个句子的结构相似，并且在关键词 *of* 上都出现句法违反:

句(1): Although Erica kissed Mary's mother, she did not kiss Dana's *of* the bride.

句(2): Although the bridesmaid kissed Mary, she did not kiss Dana's *of* the bride.

在句(1)中，Dana's 后面既可以直接结束也可以出现名词，为弱句法预期条件。与之不同，在句(2)中，Dana's 后面必须出现名词才符合句法规则，因此是强句法预期条件。结果发现，在关键词 *of* 上，强预期条件相比弱预期条件诱发了波幅更大的 ELAN，说明语言理解过程受到了句法预期的影响。另外一些研究者考察了关键词之前的加工阶段，发现此时人们已经能够预先激活句法结构(Arai & Keller, 2013; Kamide, 2012; Kuperberg & Jaeger, 2016)，比如在阅读和口语加工中都能够预先激活句法性别信息(Otten, Nieuwland, & van Berkum, 2007; Otten & van Berkum, 2008; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005)，从而更加明确地证实了句法预期的存在。

1.2 音乐结构对句法加工的影响

音乐句法加工总是发生在一定的音乐结构之中，那么句法加工如何受到音乐结构的调节？少量研究对这个问题进行过初步探讨，但是结论并不一致。

1.2.1 不同层级结构水平上的音乐句法加工

音乐在呈现过程中并不是永不停歇地进行，而是通过不同的句读标志划分出不同级别的句法结构单位。这些结构单位大致包括动机、乐节、乐句、乐段及其与它们同级的一些结构形式。根据句法规则，小的结构相互结合形成时间尺度更大的结构，大结构再结合起来，最终形成音乐层级结构(Lerdahl & Jackendoff, 1983; Zhang, Jiang, Zhou, & Yang, 2016)。音乐作品借助音乐层级结构的排列组合构成一个完整的结构系统，陈述音乐内容，表达音乐形象。

如 1.1 所述，现有音乐句法研究的实验材料通常包含少数几个和弦或音符，因此主要考察了较低层级的句法加工过程，对应于乐句或者乐句以下的结构。除了低层级的句法加工，听众是否能够将新出现的音符或和弦整合到更高层级的音乐结构之中？早期行为研究得到了否定的结果。例如，实验中将音乐片段按照小节、乐节和乐句截开再随机排列，考察这种强烈的违反是否会干扰听众对于音乐的加工。结果发现，只有在低层级结构水平上(小节和乐节)随机排列，被试对目标刺激的探测才会受到影响(Tillmann & Bigand, 1998)。同样，只有在局部结构上随机排列，人们

才能敏感地觉察出违反，从而影响他们对于音乐审美和连贯性的判断(Eitan & Granot, 2008; Granot & Jacoby, 2011)。

难道高层级的音乐结构只存在于音乐理论之中，并不具有知觉现实性？最近几项研究采用认知神经科学的方法，考察了更高层级的句法加工过程(Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013; Ma, Ding, Tao, & Yang, 2018a, 2018b)。在这些研究中，实验材料由两个乐句组成，这两个乐句构成一个乐段。因此，最后的目标和弦与第一个乐句之间的句法依存关系发生在更高层级的乐段层面。对第一个乐句进行操纵，使得目标和弦与第一个乐句之间的句法依存关系合理或者不合理。结果发现，与低层级句法违反一样，乐段结构中的句法违反同样诱发了早期的 ERAN 成分和晚期的 N5 成分，表明更高层级的音乐结构也会影响当前信息的加工。不同的是，Koelsch 等人(2013)的实验中，无论音乐家还是非音乐家都能进行高层级的音乐句法加工，但是在 Ma 等人(2018a, 2018b)的研究中，高层级的句法加工却依赖于专业的音乐训练。实验结果的差异，可能是因为实验中选用的音乐材料都是西方音乐，前一个研究是西方被试加工西方音乐，而后面的研究是中国被试加工西方音乐，此时类似于二语加工，因而难度更大。

如果在音乐加工中听众不仅能够进行低层级句法加工，也能进行更高层级的句法加工，那么不同层级结构中的句法加工是如何共同发生的？一项脑成像研究分别将音乐片段按照小节、乐句、乐段截开再随机排列，结果发现，听众能够进行不同层级的句法加工，层级越高，加工脑区从初级感觉皮层向更高级的脑区移动，在脑区上也表现出层级分布(Farbood, Heeger, Marcus, Hasson, & Lerner, 2015)。除了加工脑区，不同层级的句法加工在加工进程和加工难度等方面是否有差异？Zhang, Zhou, Chang 和 Yang (2018)使用两个乐句长度的音乐材料，采用 EEG 技术，同时考察了目标和弦与第一个乐句(远距离)和第二个乐句(近距离)的整合。结果发现，在早期阶段，只有近距离句法违反诱发 ERAN 效应，表明近距离的句法加工更早发生；在晚期阶段，近距离、远距离句法违反均诱发 N5 效应，并且不同距离的句法加工在 N5 上存在交互作用，表明不同距离的句法加工均能发生，并且低层级句法加工可能存在优先性。这项研究比较了不同距离的句法加工过程，但是由于两个乐句之间不构成层级关系，因此并没有涉及不同层级结构中的句法整合过程。最新一项研究对这个问题进行了初次探讨，研究者同时比较了听众对近距离句法依存和长距离句法依存的加工，这两种句法加工过程分别发生在乐句和乐段结构之中(Zhou, Liu, Jiang, Jiang, & Jiang, 2019)。结果发现，对于控制组被试而言，乐句内的句法违反诱发了 ERAN、N5 和 LPC 成分，乐段层面的句法违反诱发了类 ERAN 成分和 N5 成分。相关分析发现 ERAN 和类 ERAN 之间并不存在相关，可能表明高层级句法违反的探测需要额外的加工过程。对于先天性失歌症被试而言，两种层级的句法整合过程均难以进行。

1.2.2　音乐时间结构对音高句法加工的影响

音乐中的句法规则包括音高和时间两个层面(Fitch, 2013; Sun, Liu, Zhou, & Jiang, 2018)，本项目中的音乐句法指的是建立在音高结构上的组织规则。除了音高结构，构成音乐的另一个重要的结构是时间结构(张晶晶, 杨玉芳, 2017)。音乐的时间结构主要包括时长(duration)、节奏(rhythm)和节拍(meter)。时长是每个声音或者休止持续的时间，而节奏是指这些声音或休止在时间上如何组织，节拍则是拍子按照特定的强弱规律循环往复(江俊, 王梓梦, 万璇, 蒋存梅, 2014)。我们听音乐时，只有将音高结构和时间结构结合起来，才能形成连贯的音乐表征(Lerdahl & Jackendoff, 1983)。那么，音乐音高句法加工是否会受到时间结构的影响？关于这个问题存在两种理论模型，分别是单一成分模型(Jones & Boltz, 1989)和双成分模型(Peretz & Coltheart, 2003; Peretz & Kolinsky, 1993)。

单一成分模型(single-component model)预期音乐句法结构和时间结构之间存在交互作用(Jones & Boltz, 1989)。该模型认为，我们大脑中的神经元会同时追踪音高和时间维度(pitch-time entrainment)，形成联合预期，共同指导听众注意资源的分配。当音高事件出现在预期时间点，此时注意处于峰值，音高事件能够得到更好的加工。单一成分模型得到了一些实验研究的支持。例如一项行为研究发现，当音符出现的时间符合规则时，相比提前或者延迟出现，被试对音符音

高的判断更加准确(Jones, Moynihan, MacKenzie, & Puente, 2002)。另外一些研究给被试呈现一段旋律或者和弦序列，分别操纵音乐序列的句法合理性和时间合理性，要求被试对这段音乐序列的句法合理性进行判断。结果发现，被试对句法合理性的判断受到音乐时间结构的影响，当旋律或者和弦序列的时间结构违反时，句法合理性的评分显著更低(Schmuckler & Boltz, 1994; Tillmann & Lebrun-Guillaud, 2006)。除此之外，一项 fMRI 实验也显示，当被试演奏音乐序列时，音高维度和时间维度都会激活感觉运动脑区，表明两者在加工脑区上存在重叠(Brown et al., 2013)。

双成分模型(two-component mode)认为音高维度和时间维度是组成音乐的两个平行、独立的子系统，因此，音高句法的加工不受时间结构的影响(Peretz & Coltheart, 2003; Peretz & Kolinsky, 1993)。这个模型的提出首先基于神经心理学的证据。对于脑损伤患者的案例研究发现，一些患者音高加工受损，但是节奏加工与控制组相比不存在显著差异；另外一些患者无法准确地感知音乐节奏，但是音高知觉正常(Peretz, 1990, 1996)。这种临床上双分离现象提示音高和时间两个维度是独立加工的。Lebrun-Guillaud, Tillmann 和 Justus (2008)给小脑损伤患者呈现 8 和弦序列，对和弦序列终止式以及和弦序列的时间结构进行操纵，构建音高句法合理、不合理以及时间结构合理、不合理的和弦序列。结果发现，小脑损伤只对音乐时间结构的判断产生影响，并不影响音乐音高句法的加工。除了神经心理学的证据，来自正常被试的研究同样给双成分模型提供了实验支持。Palmer 和 Krumhansl (1987)给被试呈现不同版本的音乐序列：第一种是原始序列，音乐旋律中包含完整的时间结构和音高结构；对这个序列分别进行修改，使得修改版本中只保留时间结构或者音高结构。被试的任务是对音乐序列的完整度进行评分，结果发现原始版本的完整度评分是时间版本和音高版本完整度评分的线性相加，研究者据此认为音乐时间结构和音高句法结构是独立加工的。一项脑成像研究采用类似的实验材料，发现被试在演奏这些不同版本的音乐序列时，分别激活了不同的脑区，时间结构激活了腹侧视觉通路，而句法结构激活了背侧视觉通路(Bengtsson & Ullén, 2006)。

2 问题的提出

综上所述，音乐句法研究在近些年取得了较大进展，研究者们发现听众由于专业的音乐训练或者长期的内隐学习，具备了加工音乐句法的能力，并且找出了反映音乐句法加工的特定 ERP 成分以及加工脑区。在此基础之上，仍有一些重要的科学问题尚待考察。

第一，音乐句法加工的认知机制是基于整合还是预期？已有大量研究发现，当目标刺激与音乐背景之间句法不合理时，相比于句法合理，被试对于目标刺激的加工在行为和认知神经指标中都表现出更大的困难。这种加工困难存在两种可能的认知机制。第一种是整合更加困难，当目标刺激与音乐背景句法不合理时，在目标刺激出现之后，听众需要花费更多的认知资源，将目标刺激整合到先前音乐背景的表征之中。第二种是预期带来的代价，随着音乐的展开，听众根据句法知识对接下来将要出现的音乐事件建立了预期，当目标刺激与预期不一致时，听众感到诧异，导致对目标刺激加工更困难。如果要试图分离这两种可能的认知机制，研究者需要清楚地揭示在目标刺激出现之前，听众是否已经根据背景信息建立预期，预先激活了目标刺激。如果存在预先激活，则清晰地揭示了预期在句法加工中的作用；如果不存在，那么研究证据则更加支持整合观点。现有行为研究将关注点放在目标刺激出现之后，因而难以对这两种机制进行区分。认知神经科学的研究将关注点放在目标刺激出现之时，例如 EEG 研究中，句法不合理的目标刺激在 100~350 ms 出现了 ERAN 成分，尽管这种早期成分被认为反映自动或半自动化的句法违反探测过程，但是仍然难以说明这种探测过程主要是基于何种机制产生的。本项目拟深入考察目标刺激出现之前的预期阶段和出现之后的整合阶段，揭示预期和整合在音乐句法加工中的作用，以及两种机制之间可能存在的相互影响。

第二，音乐句法加工如何受到音乐结构的调节？音乐句法加工总是发生在一定的音乐结构之中，但是音乐句法加工如何受到音乐结构的调节目前并不清楚。具体而言，音乐层级结构在音乐理论中占据了重要的地位(Lerdahl & Jackendoff, 1983; Meyer, 2008)，也是作曲家组织作品、表演

家演绎作品的重要手段。但是，是否存在高层级句法加工过程仍然有待更多实验证据的支持。更重要的问题是，不同层级的句法加工过程如何共同发生几乎无人涉及，亟需对此展开深入研究。对于音乐时间结构，目前有少量研究考察了时间维度对音高句法加工的影响，但是结论并不一致。一部分研究支持独立加工的双成分模型，另一部分研究支持反映相互影响的单一成分模型，基于此，Peretz 等提出多重加工阶段理论(multiple stage processing model)，认为音高维度和时间维度在早期独立加工，在晚期阶段两者存在交互作用(Peretz & Coltheart, 2003)。那么，两个维度何时独立加工，何时会发生交互作用？对于上述两个关键问题，多重加工阶段理论仍然局限在概念框架模型对认知机制的描述阶段，需要实验证据的支持。采用具有高时间分辨率的 EEG，将有助于对这个问题进行解析。综上，本项目拟采用 EEG 技术，考察不同音乐层级结构和时间结构中音高句法的加工。

3　研究构想

3.1　研究问题和方法

本项目拟深入研究音乐句法的加工，重点解决两个问题：音乐句法加工的认知机制；音乐句法加工如何受到音乐结构的调节(图 1)。研究者拟设计 3 个系列研究考察上述两个关键问题。研究一考察预期和整合在音乐句法加工中的作用，研究二考察不同音乐层级结构中句法的加工，研究三考察音乐时间结构对音乐音高句法加工的影响。此外，音乐训练是音乐认知研究中的重要变量(Carey et al., 2015; Du & Zatorre, 2017; Nan et al., 2018; Sun et al., 2018)。在音乐句法加工中，音乐家相比非音乐家对句法违反的探测更加敏感(Koelsch, Schmidt, & Kansok, 2002; Müller, Höfel, Brattico, & Jacobsen, 2010)，表现在诱发出波幅更大的早期右前负波 ERAN。其中的原因，可能是由于在音乐呈现过程中，音乐家根据已有知识对即将出现的音乐事件建立了更强的期待(Koelsch et al., 2007)。鉴于此，本项目中 3 个研究都将考察音乐训练对实验效应的调节作用，以期更加系统深入地探讨音乐句法的加工过程。

研究工具方面，由于本项目的几个研究都需要精确考察在线加工的时间进程，因此采用具有高时间分辨率的 EEG 技术作为研究工具。对采集到的 EEG 数据分别进行 ERP 分析和时频分析。ERP 分析对 EEG 信号在时域上进行叠加平均，能够有效考察与事件锁时锁相的脑电位变化。时频分析是从时域和频域角度分析 EEG 信号，考察单个电极点上各频段的能量变化，能够观察与事件具有锁时关系的频域特征。通过两种分析，以期

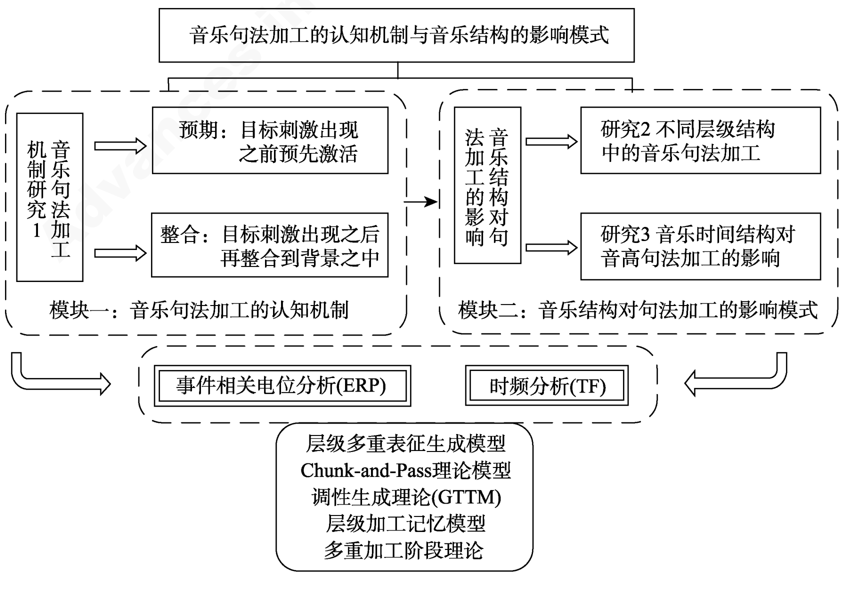


图 1　理论研究框架

得到汇聚性结果，增加实验结论的可靠性。

实验被试方面，为了考察音乐训练的调节作用，所有实验均采用两组被试，分别是受过 10 年以上专业音乐训练的音乐家，以及未接受过专业音乐训练的非音乐家。同时，对于两组被试的年龄、性别、利手、智力水平以及受教育年限等因素进行匹配。

实验材料方面，为了排除音乐熟悉性对被试加工过程的影响，所有实验均采用不熟悉的音乐作为实验材料。为了保证音乐材料符合实验设计的要求，在正式实验之前将对实验材料进行测评。

3.2 实验设计与预期结果

研究 1：音乐句法加工的认知机制：基于整合还是预期

现有研究发现听众普遍具有感知音乐句法的能力，但是其中的加工机制主要是基于整合还是预期却并不清楚。研究者们都将关注点放在新信息出现之时或者之后，因而难以对这两种机制进行分离。研究 1 拟操纵目标刺激在音乐序列中的句法合理性，以及音乐背景的预期性，同时考察目标刺激出现之前的预期阶段和出现之后的整合阶段，从而探究预期和整合在音乐句法加工中的作用。此外，选取音乐家和非音乐家两组被试，对比他们在实验效应上的差异，考察音乐训练是否影响句法加工机制。

基于音乐句法的现有研究(e.g., Koelsch et al., 2000, 2007)以及语言加工中的预期研究(e.g., DeLong et al., 2005)，本研究主要预期包括：(1)如果支持预期机制，那么在预期阶段，高预期条件和低预期条件在 ERP 波幅和时频频段上会表现出显著差异；反之，不存在任何差异；(2)在整合阶段，句法不合理音符相比句法合理音符，ERP 上诱发波幅更大的 ERAN 和 N5，时频上可能呈现早期 alpha 频段能量降低，晚期 gamma 能量降低(Ruiz, Koelsch, & Bhattacharya, 2009)；(3)音乐训练会调节句法加工的策略或者敏感程度，表现为 ERP 成分和时频频段上质的差异或者量的变化。

研究 2：不同结构层级水平上的音乐句法加工

层级加工记忆模型认为，不同时间尺度的背景信息会持续影响当前信息的加工(Hasson, Chen, & Honey, 2015)。那么，音乐句法加工如何受到不同层级结构中背景信息的影响？现有研究主要集中于考察低层级句法加工过程，更高层级结构的作用并不清楚。更重要的是，不同层级的句法加工过程如何共同发生鲜有涉及。因此，研究 2 拟选用时间尺度较长的音乐材料，包含两个不同层级的音乐结构，分别操纵目标刺激与低层级、高层级音乐背景之间的句法依存关系，形成句法合理或者不合理的音乐序列，并且选取音乐家和非音乐家两组被试，从而考察：(1)是否存在高层级句法加工过程；(2)如果存在高层级句法加工，那么不同层级的句法加工在加工进程、加工难度等方面是否存在差异；(3)对比音乐家和非音乐家两组被试在实验效应上的差异，考察音乐层级句法加工能力是否依赖于音乐训练。

基于层级加工记忆模型(Hasson et al., 2015)以及先前的实验研究(e.g., Koelsch et al., 2013; Zhang et al., 2018)，研究 2 主要包括如下预期：(1)低层级句法不合理条件相比合理条件，在 ERP 上诱发波幅更大的 ERAN 和 N5，时频上早期 alpha 频段能量降低，晚期 gamma 能量降低(Ruiz et al., 2009)；(2)高层级句法不合理会诱发与低层级不合理类似的脑电成分；(3)不同层级的句法加工在晚期时间窗口存在交互作用；(4)不同音乐层级结构中的句法加工受到音乐训练的调节，特别是高层级句法加工依赖音乐训练。

研究 3：音乐时间结构对音乐音高句法加工的影响

本项目中的句法加工是基于音乐音高结构的加工，除了音高结构，音乐中另外一个重要的维度是时间结构。只有将两个维度组合起来，才能对音乐事件形成连贯的表征(Lerdahl & Jackendoff, 1983)。音乐音高句法加工是否受到音乐时间结构的影响？现有研究对于这个问题存在很大忽视，仅有的研究也得到了冲突的结果(Tillmann & Lebrun-Guillaud, 2006)。研究 3 拟选用和弦序列，通过对终止和弦或者先前的背景和弦进行操纵，形成句法合理、不合理以及时间结构合理、不合理的和弦序列，此外选取音乐家和非音乐家两组被试，考察音乐音高句法加工是否受到音乐时间结构的影响；如果时间结构影响音高句法加工，这种影响何时以及如何发生；对比两组音乐经验不同的被试所诱发的脑电效应，考察实验效应是否受到音乐训练的调节。

基于多重加工阶段理论(Peretz & Coltheart, 2003)以及音乐句法、时间结构的现有研究，研究

3 主要包括如下预期: (1)存在句法合理性和时间结构合理性的主效应。句法违反将在 ERP 上诱发波幅更大的 ERAN 和 N5, 时频上早期 alpha 频段能量降低, 晚期 gamma 能量降低(Ruiz et al., 2009)。时间结构违反, 特别是目标和弦提前出现(Geiser, Ziegler, Jancke, & Meyer, 2009; Zhang, Che, & Yang, 2019), 会诱发波幅更大的 N150 或者 MMN; (2)音高句法加工和时间结构存在交互作用, 音乐时间结构会在早期阶段或者晚期阶段影响音高句法效应; (3)时间结构对音高句法加工的调节, 在音乐家被试中更加明显, 表现为交互作用的时间点提前, 或者实验效应更加显著。

4　理论意义

本项目在现有音乐句法研究的基础之上, 拟深入考察音乐句法加工的认知机制与音乐结构对句法加工的影响模式, 将进一步揭示音乐句法加工的本质, 对音乐理论进行检验和推进。此外, 研究结果与语言研究相互比较, 将进一步推动音乐和语言的比较研究, 对相关语言理论和更一般认知模型的验证与拓展起到积极作用。

在语言加工领域, 层级多重表征生成模型(hierarchical multi-representational generative framework, Kuperberg & Jaeger, 2016)以及 Chunk-and-Pass 模型(Christiansen & Chater, 2016)都认为语言理解是一个即时的迅速的过程, 预期在其中扮演了关键的作用。在音乐加工中, 一些音乐理论例如 ITPRA 理论(Huron, 2006)和旋律预期模型(Margulis, 2005)同样强调了预期的作用。那么, 音乐句法的加工是否也是一个即时迅速的过程, 甚至存在预先激活的预期过程? 现有音乐句法的研究都将关注点放在新信息出现之时或者之后, 因而难以对整合和预期机制进行区分。研究一拟深入考察目标刺激出现之前的预期阶段和出现之后的整合阶段, 进一步揭示音乐句法加工的认知机制, 为音乐认知模型的构建提供实证依据。同时为音乐和语言的比较研究提供新的角度, 推进人类更一般交流机制模型的生成。

音乐层级结构在音乐理论中占据了重要的地位。例如在 Lerdah 和 Jackendoff (1983)提出的著名的调性生成理论(a generative theory of tonal music, GTTM)中, 其核心观点就是认为音乐的句法结构、韵律结构等都是按照层级关系组织的。那么, 音乐层级结构是否具有认知现实性? 此外, 近些年研究者采用 fMRI 技术, 分别考察了语言(Lerner, Honey, Silbert, & Hasson, 2011)和音乐(Farbo od et al., 2015)中的层级加工过程, 发现层级越高, 加工脑区从感觉皮层向高级皮层移动, 在脑区上也表现出层级分布。在这一系列实验基础之上, Hasson 等人(2015)提出了层级加工记忆模型(hierarchical process memory model), 认为不同层级的背景信息会持续影响当前信息的加工。那么不同层级结构中的音乐背景是否会影响音乐句法的加工? 如果存在影响, 那么除了加工脑区之外, 不同层级的句法加工在加工进程和加工难度上又存在何种差异? 研究二考察不同结构层级水平上的音乐句法加工, 将对 GTTM 模型展开检验, 并且从不同层面对层级加工记忆模型进行丰富和拓展。

本项目中的音乐句法是建立在音高维度上的一组规则, 除了音高之外, 时间是音乐中另外一个重要维度, 听者需要将两者整合起来, 才能形成连贯的音乐表征。特征整合是认知心理学中的一个基本问题(Treisman & Gelade, 1980)。在音乐领域, Peretz 和 Coltheart (2003)提出多重加工阶段理论(multiple stage processing model), 认为音高维度和时间维度在早期独立加工, 在晚期阶段存在交互作用。在认知神经层面, Friston 和 Buzsáki (2016)基于语言等领域的研究, 提出了双通路模型, 认为类似于 what 和 where 通路, 在大脑中同样存在两条不同的通路, 来分别加工 what 和 when 信息。研究三采用 EEG 考察音乐时间结构对音高句法加工的影响, 可以有效检验多重加工阶段理论的心理现实性, 并且明确各加工阶段的时间进程信息。同时以音乐材料为载体, 从电生理的角度进一步考察 what (音高)和 when (时间)加工的认知神经机制, 推动一般认知模型的构建。

参考文献

江俊, 王梓梦, 万璇, 蒋存梅. (2014). 音乐时间加工的影响因素. *心理科学进展, 22*(4), 650–658.

马谐, 杨玉芳, 张秋月. (2016). 音乐句法的加工. *科学通报, 61*(10), 1099–1111.

叶铮, 周晓林. (2006). 音乐之脑. *心理科学进展, 14*(5), 641–647.

张晶晶, 杨玉芳. (2017). 音乐句法加工的影响因素. *心理科学进展, 25*(11), 1823–1830.

周临舒, 蒋存梅, 杨玉芳. (2012). 音乐和语言句法认知的比较. *科学通报, 57*(28), 2674–2685.

Arai, M., & Keller, F. (2013). The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes, 28*(4), 525–560.

Bengtsson, S. L., & Ullén, F. (2006). Dissociation between melodic and rhythmic processing during piano performance from musical scores. *NeuroImage, 30*(1), 272–284.

Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception & Psychophysics, 41*(6), 519–524.

Bigand, E., & Pineau, M. (1997). Global context effects on musical expectancy. *Perception & Psychophysics, 59*(7), 1098–1107.

Bigand, E., Tillmann, B., Poulin, B., D'Adamo, D. A., & Madurell, F. (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition, 81*(1), B11–B20.

Brown, R. M., Chen, J. L., Hollinger, A., Penhune, V. B., Palmer, C., & Zatorre, R. J. (2013). Repetition suppression in auditory-motor regions to pitch and temporal structure in music. *Journal of Cognitive Neuroscience, 25*(2), 313–328.

Carey, D., Rosen, S., Krishnan, S., Pearce, M. T., Shepherd, A., Aydelott, J., & Dick, F. (2015). Generality and specificity in the effects of musical expertise on perception and cognition. *Cognition, 137*, 81–105.

Chen, Q., Zhang, J., Xu, X., Scheepers, C., Yang, Y., & Tanenhaus, M. K. (2016). Prosodic expectations in silent reading: ERP evidence from rhyme scheme and semantic congruence in classic Chinese poems. *Cognition, 154*, 11–21.

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences, 39*, 1–72.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117–1121.

Du, Y., & Zatorre, R. J. (2017). Musical training sharpens and bonds ears and tongue to hear speech better. *Proceedings of the National Academy of Sciences, 114*(51), 13579–13584.

Eitan, Z., & Granot, R. Y. (2008). Growing oranges on Mozart's apple tree: "Inner form" and aesthetic judgment. *Music Perception, 25*(5), 397–418.

Farbood, M. M., Heeger, D. J., Marcus, G., Hasson, U., & Lerner, Y. (2015). The neural processing of hierarchical structure in music and speech at different timescales. *Frontiers in Neuroscience, 9*, 157.

Fitch, W. T. (2013). Rhythmic cognition in humans and animals: Distinguishing meter and pulse perception. *Frontiers in systems neuroscience, 7*, 68.

Friston, K., & Buzsáki, G. (2016). The functional anatomy of time: What and when in the brain. *Trends in Cognitive Sciences, 20*(7), 500–511.

Geiser, E., Ziegler, E., Jancke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex, 45*(1), 93–102.

Granot, R. Y., & Jacoby, N. (2011). Musically puzzling I: Sensitivity to overall structure in the sonata form? *Musicae Scientiae, 15*(3), 365–386.

Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: Memory as an integral component of information processing. *Trends in Cognitive Sciences, 19*(6), 304–313.

Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT press.

Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language, 86*, 157–171.

Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language, 98*, 1–11.

Jentschke, S., Friederici, A. D., & Koelsch, S. (2014). Neural correlates of music-syntactic processing in two-year old children. *Developmental Cognitive Neuroscience, 9*, 200–208.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*(3), 459–491.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13*(4), 313–319.

Kamide, Y. (2012). Learning individual talkers' structural preferences. *Cognition, 124*(1), 66–71.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163–182.

Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience, 15*(3), 170–180.

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: "nonmusicians" are musical. *Journal of Cognitive Neuroscience, 12*(3), 520–541.

Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology, 44*(3),

476–490.

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences, 110*(38), 15443–15448.

Koelsch, S., Schmidt, B.-H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology, 39*(5), 657–663.

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences, 23*(1), 63–77.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31*(1), 32–59.

Lagrois, M.-É., Peretz, I., & Zendel, B. R. (2018). Neurophysiological and behavioral differences between older and younger adults when processing violations of tonal structure in music. *Frontiers in Neuroscience, 12*, 54.

Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and language, 98*(1), 74–88.

Lebrun-Guillaud, G., Tillmann, B., & Justus, T. (2008). Perception of tonal and temporal structures in chord sequences by patients with cerebellar damage. *Music Perception, 25*(4), 271–283.

Lerdahl, F., & Jackendoff, R. S. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT press.

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *Journal of Neuroscience, 31*(8), 2906–2915.

Li, X., Zhang, Y., Xia, J., & Swaab, T. Y. (2017). Internal mechanisms underlying anticipatory language processing: Evidence from event-related-potentials and neural oscillations. *Neuropsychologia, 102*, 70–81.

Ma, X., Ding, N., Tao, Y., & Yang, Y. F. (2018a). Differences in neurocognitive mechanisms underlying the processing of center-embedded and non-embedded musical structures. *Frontiers in Human Neuroscience, 12*, 425.

Ma, X., Ding, N., Tao, Y., & Yang, Y. F. (2018b). Syntactic complexity and musical proficiency modulate neural processing of non-native music. *Neuropsychologia, 121*, 164–174.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience, 4*(5), 540–545.

Maess, B., Mamashli, F., Obleser, J., Helle, L., & Friederici, A. D. (2016). Prediction signatures in the brain: Semantic pre-activation during language comprehension. *Frontiers in Human Neuroscience, 10*, 591.

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception, 22*(4), 663–714.

Meyer, L. B. (2008). *Emotion and meaning in music*. Chicago, IL: University of chicago Press.

Müller, M., Höfel, L., Brattico, E., & Jacobsen, T. (2010). Aesthetic judgments of music in experts and laypersons—An ERP study. *International Journal of Psychophysiology, 76*(1), 40–51.

Nan, Y., Liu, L., Geiser, E., Shu, H., Gong, C. C., Dong, Q., ... & Desimone, R. (2018). Piano training enhances the neural processing of pitch and improves speech perception in Mandarin-speaking children. *Proceedings of the National Academy of Sciences, 115*(28), E6630–E6639.

Otten, M., Nieuwland, M. S., & van Berkum, J. J. (2007). Great expectations: Specific lexical anticipation influences the processing of spoken language. *BMC neuroscience, 8*(1), 89.

Otten, M., & van Berkum, J. J. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes, 45*(6), 464–496.

Palmer, C., & Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13*(1), 116–126.

Patel, A. D. (2010). *Music, language, and the brain.* Oxford: Oxford university press.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10*(6), 717–733.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain, 113*(4), 1185–1205.

Peretz, I. (1996). Can we lose memory for music? A case of music agnosia in a nonmusician. *Journal of Cognitive Neuroscience, 8*(6), 481–496.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*(7), 688–691.

Peretz, I., & Kolinsky, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: A neuropsychological perspective. *The Quarterly Journal of Experimental Psychology, 46*(2), 301–325.

Poulin-Charronnat, B., Bigand, E., & Koelsch, S. (2006). Processing of musical syntax tonic versus subdominant: An event-related potential study. *Journal of Cognitive Neuroscience, 18*(9), 1545–1554.

Ruiz, M. H., Koelsch, S., & Bhattacharya, J. (2009). Decrease in early right alpha band phase synchronization and late

gamma band oscillations in processing syntax in music. *Human Brain Mapping, 30*(4), 1207–1225.

Schmuckler, M. A., & Boltz, M. G. (1994). Harmonic and rhythmic influences on musical expectancy. *Perception & Psychophysics, 56*(3), 313–325.

Sun, L., Liu, F., Zhou, L., & Jiang, C. (2018). Musical training modulates the early but not the late stage of rhythmic syntactic processing. *Psychophysiology, 55*(2), e12983.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632–1634.

Tillmann, B., & Bigand, E. (1998). Influence of global structure on musical target detection and recognition. *International Journal of Psychology, 33*(2), 107–122.

Tillmann, B., Bigand, E., & Pineau, M. (1998). Effects of global and local contexts on harmonic expectancy. *Music Perception, 16*(1), 99–117.

Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research, 16*(2), 145–161.

Tillmann, B., & Lebrun-Guillaud, G. (2006). Influence of tonal and temporal expectations on chord processing and on completion judgments of chord sequences. *Psychological Research, 70*(5), 345–358.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136.

van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(3), 443–467.

van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology, 83*(2), 176–190.

Zhang, J., Che, X., & Yang, Y. (2019). Event-related brain potentials suggest a late interaction of pitch and time in music perception. *Neuropsychologia, 132*, 107118.

Zhang, J., Jiang, C., Zhou, L., & Yang, Y. (2016). Perception of hierarchical boundaries in music and its modulation by expertise. *Neuropsychologia, 91*, 490–498.

Zhang, J., Zhou, X., Chang, R., & Yang, Y. (2018). Effects of global and local contexts on chord processing: An ERP study. *Neuropsychologia, 109*, 149–154.

Zhou, L., Liu, F., Jiang, J., Jiang, H., & Jiang, C. (2019). Abnormal neural responses to harmonic syntactic structures in congenital amusia. *Psychophysiology*, e13394.

The cognitive mechanism of music syntactic processing and the influence of music structure on its processing

ZHANG Jingjing[1]; LIANG Xiaoyue[2]; CHEN Yidi[2]; CHEN Qingrong[1]

([1] *School of Psychology, Nanjing Normal University, Nanjing* 210097, *China*)

([2] *Conservatory of Music, Nanjing Normal University, Nanjing* 210097, *China*)

Abstract: Music and language are the two most important sign systems for human beings. Similar to language, music is also constructed on certain syntactic rules. Although evidence has shown that listeners are sensitive to musical syntax, the underlying cognitive mechanism and its influencing factors are still unknown. Therefore, we intend to further explore the role of prediction and integration in musical syntactic processing, as well as the impact of musical hierarchical and temporal structures on musical syntactic processing. We expect the proposed studies to further reveal the nature of musical syntactic processing, and to provide empirical evidence for the comparison between music and language and the exploration of human general communication mechanism.

Key words: musical syntax; hierarchical structures; temporal structures; cognitive mechanism; neural mechanisms